14

Proving Causation in Damages Analyses

ONCE a defendant has been accused of engaging in some unlawful act—securities fraud, a breach of contract, or participation in a price-fixing conspiracy—there will come a time when an economist will be asked to compute the amount of damages that should be awarded to the plaintiff should liability be proven. Assessing damages is a complex undertaking because in many cases, the legal standard requires an estimate of the injury that would not have happened but for the unlawful act: that is, the assessed injury must have been caused by the unlawful acts.

This chapter describes the challenges in determining the damages that can be causally connected to the unlawful act that led to them. Although econometric and other statistical methods are often used to isolate the impact of the unlawful act from other events in the marketplace, an empirical correlation between the "bad act" and the calculated damages does not imply causation. Conclusions regarding causality must therefore be based on economic theory or some a priori chain of reasoning capable of explaining why and how the cause led to the effect.

Proving Causation in Damages Analyses

Lauren Stiroh

One of my colleagues in graduate school kept a newspaper clipping with the title "Correlation Isn't Causation, But Who Cares?" taped to his door. The content of the original article is lost in the passage of time, but the sentiment is as relevant today as it ever was. The distinction between correlation and causation is often willfully blurred by experts or lawyers proving damages, most likely in the hope that having proven the "fact of damage" the court will allow a plaintiff substantially less rigor in proving the amount of damages.

The legal standard for determining the amount of damages in a variety of cases (commercial litigation, securities fraud, breach of contract, etc.) is but-for causation: "the injury to an individual would not have happened but for the conduct of the wrongdoer."[1] The but-for standard implies that the amount of damages be causally connected to the acts that led to them. In spite of this requirement, damages experts are often satisfied with showing the amount of damages that are associated with, but were not necessarily caused by, the complained-of actions. In this chapter, I describe what causation means to an economist and the econometric tools that can be brought to bear to prove or disprove causation in damages cases.

Correlation versus Causation

Correlation is the occurrence of events in a predictably consistent sequence. Causation implies that if a first event had not occurred, a second event would not have happened (e.g., if the defendant had not infringed its rival's patent, the defendant would not have captured its rival's customers). Events can be correlated without being causally connected.

For example, the outcome of the National Football League's Super Bowl is often viewed as a predictor of activity in the US stock market. Sports enthusiasts (and people with time on their hands) have noted that the stock market tends to fall in a year when a team from the old American Football League (AFL) wins the Super Bowl. In fact, since 1967, the Super Bowl indicator has accurately predicted gains and slumps in the stock market

about 80 percent of the time.[2] In spite of this high degree of correlation, there is no causal link between the historical league affiliation of the winner of the Super Bowl and market activity. The victory by the New England Patriots (an old AFL team) over the Carolina Panthers in the 2004 Super Bowl did not *cause* the market to fall in 2004. There is simply no economic or other theory that suggests that but for Patriots' kicker Adam Vinatieri successfully completing a field goal in the final minute of the game, the market would have continued to rise, but, because of it, the market was destined to fall.[3] However, there is economic theory that suggests that the market will rise or fall depending on changes in the underlying economic fundamentals (productivity, technology, and availability of factors of production) and on the expectations of market participants.[4]

The distinction between correlation and causation is the presence of a theory, a chain of reasoning that explains why the cause leads to the effect. That the effect has followed the cause in the past is not sufficient. Causation implies that but for one event, the second would not happen. For example, consider the fact that the long-run correlation between the rate at which money is printed and the general inflation rate in the economy is about 80 percent.[5] Is this merely a correlation or evidence of a causal relationship? In this case, we accept that there is a causal connection between money growth and inflation because there is a theory connecting the action (printing money) to the effect (inflation). Printing more money *causes* inflation because, simply put, there are more dollars chasing after the same number of goods. Moreover, holding all else constant, if we did not print money at a faster pace, there would not be higher inflation. Without a theory or chain of reasoning linking cause to effect, correlation, no matter how strong, is a statistical curiosity.

The Use of Econometrics in Determining Causation

While accepting that correlation does not prove causation, correlation is nonetheless fundamentally important as a necessary precondition for causation to be present. In the eighteenth century, philosopher David Hume defined two events A and B as being causally connected if "events like" A were always followed by "events like" B. Around the turn of the twentieth century, Karl Pearson, a mathematician and one of the founders of statistics, added some mathematical rigor to the concept by defining causation as perfect correlation. Unfortunately for economists, perfect correlation is virtually impossible to identify. Correlation itself is difficult to measure perfectly and generally one must contend with other factors that affect the variables under measurement, muddying the waters. Even without addressing such measurement difficulties, Hume himself was skeptical that *any* amount of past experience of "events like" B following "events like" A would be sufficient to allow one to conclude that A was the cause of B.

Physicists have a distinct advantage over economists in proving causation between action and effect: a physicist can design an experiment in a vacuum such that the action under investigation is isolated from all other factors that could affect the outcome. Economists usually do not have the same luxury. In a market economy there are innumerable factors that could affect a firm's profits or sales. The damage expert's job is to isolate the impact of a specific set of actions or factors (such as an alleged price-fixing conspiracy) on profits from the impact of natural market fluctuations (such as those arising from technological progress, changes in demand, changing competitive pressures, etc.). Because the economist cannot set up an experiment that allows him to rewind the time period over which the damage was alleged to occur and replay the market events without the bad acts in question, he often relies upon statistical tools to attempt to isolate the impact of the actions under investigation from the impact of natural market forces that are not being challenged by the plaintiff.

Econometrics is a statistical tool that can be used to identify relationships between economic variables. An econometric model, properly constructed, is the reduction of a theory of interrelationships between economic variables to a single equation or a set of equations that allow economists to measure statistically the marginal effects of a set of factors on a variable of interest. Regression analysis captures how changes in one variable (the dependent variable) can be explained by changes in other variables. However, by itself, the results of an econometric analysis do not imply causation. That is because the coefficients of a regression model depend on the measured *correlation* between the variables of interest. Causation is *inferred* from a theory of how those variables interact. As the authors of one statistics textbook caution: "A statistical relationship, however strong and however suggestive, can never establish causal connection: our ideas of causation must come from outside statistics, ultimately from some theory or other."[6]

Because of this insight, economists have sought to develop ways of melding theory with measurement, looking simultaneously for the presence of patterns in the data that are consistent with causation and the absence of patterns that would disprove a causal relationship. In 1969, Clive

Granger (economist and winner of the 2003 Nobel Prize in Economics) proposed a test for causality based on how well the past can predict the future.[7] A variable X Granger-causes variable Y if predictions of Y using X are better than predictions of Y using past values of Y alone. In 1980, Christopher Sims expanded upon Granger's idea and proposed that if causation between A and B were unidirectional (i.e., A causes B but B does not cause A), then past values of A could be used to predict future values of B, but future values of A would be uncorrelated with past values of B.[8]

One difficulty with inferring causality by using historical trends to predict future values is that the method assumes that an event consistently preceding the effect is the cause. This is not necessarily the case: one can show that Christmas card sales Granger-cause Christmas, although rationality clearly tells us that causality runs in the opposite direction. A significant problem to contend with in economics is that the direction of causation is often unknown. In some cases, logic dictates which is the cause and which is the effect; for instance, few would argue that a positive correlation between the rate of crime and the size of the police force is evidence that expanding the city's police force causes more crime.

In other cases, though, it is not clear which variable is the cause and which is the effect. For example, one may observe that profit margins are higher in industries where a small number of firms account for a large proportion of sales (i.e., highly concentrated industries). However, it is not immediately obvious (and, in fact, is subject to a great deal of study and discussion) whether greater market concentration leads to higher profits or whether more efficient firms with higher profits tend to grow faster and therefore account for a greater proportion of an industry's sales. In circumstances such as these, Granger's method does not help identify the direction of causation.

The explosive growth in computing power and the availability of vast amounts of electronic data have given statisticians a greater ability to conduct more complex statistical tests to find relationships among variables. The relative ease with which one can search for relationships within the data, however, means that great care must be taken to ensure that the underlying theory of causation is sound and to rule out alternative theories of causation.

Pitfalls in Using Econometrics to Prove Causation

Some of the pitfalls in proving causation using econometric methods are highlighted in the examples above. As noted, one may encounter spurious

correlation; that is, a measurable pattern connecting the movements of two unrelated variables (e.g., the historical conference affiliation of the winner of a football game and growth in the stock market). Spurious correlation can be ruled out as a causal relationship only by a theory or lack of a theory connecting the two events. If there is no chain of reasoning that logically connects a change in X to a change in Y, then the observed correlation is dismissed as spurious.

Another problem is that the two variables in question may be strongly affected by a third variable that is the true cause of both. For example, consider the correlation between obtaining a degree from a high-ranking college and obtaining a high-paying job. Concluding that an Ivy League education causes one to obtain a high-paying job is plausible, but it ignores the role played by a person's innate ability in both gaining acceptance to an Ivy League institution *and* succeeding in one's job. If the Ivy League education is merely a signal of greater ability, then the observed correlation is not indicative of a causal relationship between an educational institution's ranking and job status.

In addition, there may be a feedback relationship between the two variables in question so that they jointly cause each other. This is referred to by economists as simultaneity or endogeneity. Consider, for example, the relationship between illiteracy and poverty. Illiteracy rates are highest among the poorest families. While one might reasonably conclude that illiteracy causes poverty because illiteracy restricts ones ability to apply for and perform high-paying jobs, one must also consider the possibility that poverty begets illiteracy because the impoverished place finding food and shelter above achieving an education.

Finally, there could be multiple causes that influence a variable of interest. For instance, a company's sales are likely to be a function of the attributes of its products, the price the company charges, the prices charged by competitors, and the company's marketing campaign, as well as those of its competitors. This problem is most easily dealt with by ensuring that the econometric model (a regression analysis, for example) includes all of the important variables that were likely to have affected the variable of interest during the damage period.

These problems highlight the need for a reasoned interpretation of the results of an econometric model. Overcoming these problems is a two-step procedure. First, one must take care that the econometric model accurately reflects reality in that it accounts for the complex relationships affecting the variables in the regression. Second, one must take care in

interpreting the results of a regression. The existence of a significant coefficient in a regression cannot be interpreted as evidence of a specific causal relationship unless (a) one can rule out or control for other causal factors and (b) there are no competing logical explanations for the results.

Determining Causation When Analyzing Damages

The legal concept of causation in damages captures the extent to which the alleged illegal conduct is the cause of injury. "Conduct is the cause of a result when...it is an antecedent but for which the result in question would not have occurred."[9] The concept also involves disaggregating the effect of unlawful from lawful factors that may have adversely affected the plaintiff. An economist who has been asked to estimate damages first identifies the but-for world (i.e., the world that the plaintiff would have experienced but for the defendant's acts). The second step is to quantify the relevant variables that describe the but-for world. Finally, the damages expert calculates the damages that the plaintiff sustained by not being able to operate in the but-for world.

The question the economist seeks to answer is, But for the acts alleged by the plaintiff, what profits would the parties to the litigation have earned? The difference between the actual profits and but-for profits is the amount of damages incurred by the plaintiff or the amount of ill-gotten gains achieved by the defendant. Note that damages need not have occurred in the past to be measurable by econometric techniques. Damages arising from securities fraud or a breach of contract, for example, may entail reduced earnings in future years. In these cases, damages can be measured by the difference in the profits that are currently *expected* given that the fraud or breach (or other objectionable act) occurred and the profits that otherwise would have been *expected* in the normal course of business.[10]

The estimation of damages is meaningful only if the economist is able to ascribe the differences between the actual and the but-for world to the alleged bad acts of the defendant. Establishing causality is therefore synonymous with designing a good damages model. A good damages model is based on a logical and well-specified theory that connects the alleged bad acts to the amount of damage.

The key attribute of a good damages model is the description of a realistic but-for world. The model must take into account the possibility that the variable in question might have changed anyway notwithstanding the alleged bad acts. To be useful for the court, the empirical model must isolate the impact of the alleged wrongdoing on the revenues and profits of the plaintiff as distinct from changes in the plaintiff's profits over the damage period that are due to the effects of lawful conduct. For example, in identifying damages associated with an allegedly false advertising campaign, the econometric damages model ideally should take into account and isolate the impact of other factors (such as a general downturn in the market, competition from a new entrant, price erosion, or rising costs) that could have affected the plaintiff's sales or profits at the time that the challenged ads were being run. Only by doing this, does the model isolate damages from changes in the plaintiff's profits that would have occurred regardless of the alleged bad acts.

Conclusion

In the eighteenth century, Hume was reluctant to draw any conclusions regarding causation, believing that one could never derive sufficient experience from past events to be secure in predicting the future from the past. His urging of "mitigated skepticism" in drawing causal inferences remains an appropriate reminder for today's damages experts and courts of law.

While econometrics is a useful tool in drawing inferences, to ascribe causality, appeal must be made to a priori or theoretical considerations. In recommending a particular theory of damages to the court, the damages expert ought to give as much consideration to illustrating the a priori chain of reasoning connecting the calculated damages amount and the unlawful action allegedly causing it as to ruling out competing potential causes of that same damages amount.

Notes

1. *Black's Law Dictionary* 6th edition (St. Paul, Minnesota: West Pub. Co., 1990); see entry for "cause-in-fact."
2. Based on annual changes in the Dow Jones Industrial Average.
3. In fact, the market did rise in 2004, albeit at a slower pace than it had in 2003, when the National Football Conference (NFC) Tampa Bay Buccaneers won the Super Bowl.
4. I note though, in the interests of full disclosure, that Mr. Vinatieri also successfully completed a field goal kick in the final quarter of the Super Bowl game in 2002, giving the Patriots a victory. Following the Patriots' 2002 victory, the market fell 14 percent over the course of the year. Superstitious readers may want to reevaluate their risk profile.

5. The economic relationship between money growth and inflation is a long-run relationship. The high correlation is observed by taking 10-year rolling averages of the two variables and comparing variations in the resulting two series. The relationship between the variables is barely perceptible on a year-to-year or month-to-month basis.

6. Maurice Kendall and Alan Stuart, *The Advanced Theory of Statistics* 4th ed., vol. 2: *Inference and Relationship* (New York, New York: Macmillan Publishing Co., 1979).

7. Clive W.J. Granger, "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica* 37 (July 1969): 424-438.

8. Christopher A. Sims, "Money, Income, and Causality," *American Economic Review* (September 1972): 540-552.

9. Model Penal Code, Section 2.03(1) (1985).

10. For a discussion of the use of event studies in determining the present value of future damages caused by fraud or a breach of contract from changes in a company's stock price concurrent with information about the fraud or breach reaching the market, see David I. Tabak and Frederick C. Dunbar, "Materiality and Magnitude: Event Studies in the Courtroom," in *Litigation Services Handbook: The Role of the Financial Expert* 3rd ed., eds. Roman L. Weil, Michael J. Wagner, and Peter B. Frank (New York, New York: J. Wiley & Sons, 2001), at 19.1.

15

Do Not-For-Profit Firms Behave Differently?

IN the area of antitrust and competition policy, the treatment of not-for-profit firms has been the subject of debate and commentary for decades. Legal discussions often focus on antitrust liability and the appropriate application of the antitrust laws to not-for-profit firms. The economic debate is equally spirited. Do not-for-profit firms behave differently? Does the conventional price-concentration paradigm hold for not-for-profit firms?

This chapter discusses the theoretical and empirical research on the issue, which, for the most part, has been analyses of ownership structure and competition in the hospital service and health insurance industries. The author finds that differences in objectives do not mean that for-profit and not-for-profit firms will make different economic decisions. One important reason is that market forces matter: in competitive markets, for-profit and not-for-profit firms tend to behave similarly. Moreover, economic theory suggests that if there are differences in behavior, they would appear in more concentrated markets. However, as the author points out, whether there is, in fact, a difference is an empirical issue that will vary from market to market.